(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
1 May 2003 (01.05.2003)

PCT

(10) International Publication Number
WO 03/036982 A2

(51) International Patent Classification[7]: H04N 7/26

(21) International Application Number: PCT/IB02/04373

(22) International Filing Date: 21 October 2002 (21.10.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
01204066.3 26 October 2001 (26.10.2001) EP
02075916.3 8 March 2002 (08.03.2002) EP

(71) Applicant *(for all designated States except US)*: **KONINKLIJKE PHILIPS ELECTRONICS N.V.** [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **BRULS, Wilhelmus, H., A.** [NL/NL]; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). **KLEIN GUNNEWIEK, Reinier, B., M.** [NL/NL]; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).

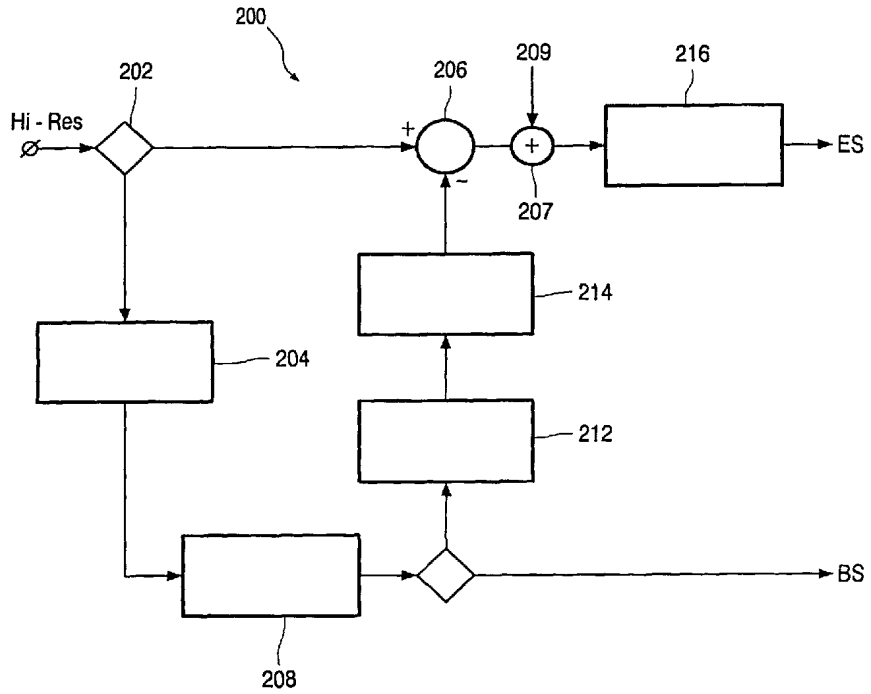
(74) Agent: **GROENENDAAL, Antonius, W., M.**; Internationaal Octrooibureau B.V., Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: CODING



(57) Abstract: The invention provides spatial scalable compression of video, comprising base layer encoding for providing a bitstream with a relatively low resolution, and enhancement layer encoding for encoding a residual signal for providing a second bitstream, wherein a modification is provided prior to the enhancement layer encoding for transforming the residual signal into a signal with a level range of a normal input video signal.

WO 03/036982 A2

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Coding

FIELD OF THE INVENTION

The invention relates to a video coding.

BACKGROUND OF THE INVENTION

Because of the massive amounts of data inherent in digital video, the transmission of full-motion, high-definition digital video signals is a significant problem in the development of high-definition television. More particularly, each digital image frame is a still image formed from an array of pixels according to the display resolution of a particular system. As a result, the amounts of raw digital information included in high resolution video sequences are massive. In order to reduce the amount of data that must be sent, compression schemes are used to compress the data. Various video compression standards or processes have been established, including, MPEG-2, MPEG-4, and H.263.

Many applications are enabled where video is available at various resolutions and/or qualities in one stream. Methods to accomplish this are loosely referred to as scalability techniques. There are three axes on which one can deploy scalability. The first is scalability on the time axis, often referred to as temporal scalability. Secondly, there is scalability on the quality axis, often referred to as signal-to-noise scalability or fine-grain scalability. The third axis is the resolution axis (number of pixels in image) often referred to as spatial scalability or layered coding. In layered coding, the bitstream is divided into two or more bitstreams, or layers. Each layer can be combined to form a single high quality signal. For example, the base layer may provide a lower quality video signal, while the enhancement layer provides additional information that can enhance the base layer image.

In particular, spatial scalability can provide compatibility between different video standards or decoder capabilities. With spatial scalability, the base layer video may have a lower resolution than the input video sequence, in which case the enhancement layer carries information which can restore the resolution of the base layer to the input sequence level.

Figure 1 illustrates a known spatial scalable video encoder 100. The depicted encoding system 100 accomplishes layer compression, whereby a portion of the channel is

used for providing a low resolution base layer and the remaining portion is used for transmitting edge enhancement information, whereby the two signals may be recombined to bring the system up to high-resolution. The high resolution video input Hi-Res is split by splitter 102 whereby the data is sent to a low pass filter 104 and a subtraction circuit 106. The low pass filter 104 reduces the resolution of the video data, which is then fed to a base encoder 108. In general, low pass filters and encoders are well known in the art and are not described in detail herein for purposes of simplicity. The encoder 108 produces a lower resolution base stream which can be broadcast, received and via a decoder, displayed as is, although the base stream does not provide a resolution which would be considered as high-definition.

The output of the encoder 108 is also fed to a decoder 112 within the system 100. From there, the decoded signal is fed into an interpolate and upsample circuit 114. In general, the interpolate and upsample circuit 114 reconstructs the filtered out resolution from the decoded video stream and provides a video data stream having the same resolution as the high-resolution input. However, because of the filtering and the losses resulting from the encoding and decoding, loss of information is present in the reconstructed stream. The loss is determined in the subtraction circuit 106 by subtracting the reconstructed high-resolution stream from the original, unmodified high-resolution stream. The output of the subtraction circuit 106 is fed to an enhancement encoder 116 which outputs a reasonable quality enhancement stream.

SUMMARY OF THE INVENTION

It is an object of the invention to provide advantageous coding of the enhancement stream. To this end, the invention provides an apparatus, a layered encoder, a method, a decoder and a method for decoding as defined in the independent claims. Advantageous embodiments are defined in the dependent claims.

It is an insight of the inventors that a prior art enhancement encoder needs non-standard components to process the residual signal which has pixel values concentrated around zero. Accordingly, there is a need for a method and apparatus which changes the concentration of pixel values of the residual signal before the residual signal is supplied to the enhancement encoder.

By providing a modification unit prior to the enhancement unit for transforming the residual signal into a signal with a signal level range of a normal input video signal, standard components can be used in the enhancement coder. Preferably, the

modification unit consists of an adder for adding a DC offset to the residual signal, preferably such that the pixel values of the residual signal are shifted to the middle of a predetermined input range of the enhancement coder.

According to one embodiment of the invention, a method and apparatus for providing spatial scalable compression using adaptive content filtering of a video stream is disclosed. The video stream is downsampled to reduce the resolution of the video stream. The downsampled video stream is encoded to produce a base stream. The bsae stream is decoded and upconverted to produce a reconstructed video stream. The reconstructed video stream is subtracted from the video stream to produce a residual stream. The residual signal is transformed into a signal with a signal level range of a normal input video signal used for video compression. The resulting residual stream is encoded and outputted as an enhancement stream.

According to another embodiment of the invention, a method and apparatus for decoding compressed video information received in a base stream and an enhancement stream is disclosed. The base stream is decoded and then upconverted to increase the resolution of the decoded base stream. The enhancement stream is decoded. A signal range of the decoded enhancement stream is transformed into a signal range of an original residual signal. The upconverted decoded base stream is combined with the transformed enhancement stream to produce a video output.

These and other aspects of the invention will be apparent from and elucidated with reference to the embodiments described hereafter.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention will now be described, by way of example, with reference to the accompanying drawings, wherein:

Figure 1 is a block schematic representation of a known encoder with spatial scalability;

Figure 2 is a block schematic representation of an encoder with spatial scalability according to one embodiment of the invention;

Figure 3 is a block schematic representation of a decoder according to one embodiment of the invention.

DETAILED DESCRIPTION OF THE INVENTION

Figure 2 illustrates a spatial scalable video encoder 200 according to one embodiment of the invention. The depicted encoding system 200 accomplishes layer compression, whereby a portion of the channel is used for providing a low resolution base layer and the remaining portion is used for transmitting edge enhancement information, whereby the two signals may be recombined to bring the system up to high-resolution. The high resolution video input is split by splitter 202 whereby the data is sent to a low pass filter 204 and a subtraction circuit 206. The low pass filter 204 reduces the resolution of the video data, which is then fed to a base encoder 208. In general, low pass filters and encoders are well known in the art and are not described in detail herein for purposes of simplicity. The encoder 208 produces a lower resolution base stream BS which can be broadcast, received and via a decoder, displayed as is, although the base stream does not provide a resolution which would be considered as high-definition.

The output of the encoder 208 is also fed to a decoder 212 within the system 200. From there, the decoded signal is fed into an interpolate and upsample circuit 214. In general, the interpolate and upsample circuit 214 reconstructs the filtered out resolution from the decoded video stream and provides a video data stream having the same resolution as the high-resolution input. However, because of the filtering and the losses resulting from the encoding and decoding, loss of information is present in the reconstructed stream. The loss is determined in the subtraction circuit 206 by subtracting the reconstructed high-resolution stream from the original, unmodified high-resolution stream. The output of the subtraction circuit 206 is fed into a modification unit 207. The modification unit 207 transforms the residual signal into a signal within the same signal level range as a normal input video signal as used for video compression. The modification unit 207 adds a DC-offset value 209 to the residual signal. The modification unit 207 also comprises a clip function which prevents the output of the modification unit from going below a predetermined value and above another predetermined value. This DC-offset and clipping operation allows the use of existing standards, e.g., MPEG, for the enhancement encoder where the pixel values are in a predetermined range, e.g., 0...255. The residual signal is normally concentrated around zero. By adding a DC-offset value 209, the concentration of samples can be shifted to the middle of the range, e.g., 128 for 8 bit video samples. The transformed residual signal from the modification unit 207 is fed to an enhancement encoder 216 which outputs a reasonable quality enhancement stream ES. The advantage of this addition is that the standard

components of the encoder for the enhancement layer can be used and result in a cost efficient (re-use of IP blocks) solution.

Figure 3 illustrates a decoder 300 according to one embodiment of the invention. Figure 3 illustrates a decoder for decoding the encoded signals processed by the layered encoder 200. The base stream is decoded in a base decoder 302. The decoded output from the decoder 302 is upconverted by an upconverter 306 and then supplied to an addition unit 310. The enhancement stream is decoded in a decoder 304. The output of the decoder 304 is supplied to a modification unit 308. The modification unit 308 performs the inverse operation of the modification 207 in the encoder 200. The modification unit 308 converts the decoded enhancement stream from a normal video signal range to the signal range of the original residual signal. The output of the modification unit 308 is supplied to the addition unit 310, where it is combined with the output of the upconverter 306 to form the output of the decoder 300.

It will be understood that the different embodiments of the invention are not limited to the exact order of the above-described steps as the timing of some steps can be interchanged without affecting the overall operation of the invention. Furthermore, the term "comprising" does not exclude other elements or steps, the terms "a" and "an" do not exclude a plurality and a single processor or other unit may fulfill the functions of several of the units or circuits recited in the claims.

CLAIMS:

1. An apparatus for performing spatial scalable compression of video informatic captured in an input signal, comprising:

a base encoder for encoding a bitstream having a relatively low resolution;

an enhancement encoder for encoding a residual signal to produce a second bitstream; and

a modification unit prior to the enhancement encoder for transforming the residual signal into a signal with a signal level range of a normal input video signal used for video compression, the residual signal being a differential signal created from a reconstructe base stream and the input signal.

2. The apparatus according to claim 1, wherein the modification unit adds a DC-offset value to the residual signal.

3. The apparatus according to claim 2, wherein the modification unit comprises clip operation which prevents an output of the modification unit from going below a first predetermined value and above a second predetermined value.

4. The apparatus according to claim 2, wherein the DC-offset value shifts a majority of pixel values of the residual signal from one end of a predetermined pixel value range to approximately a middle of the predetermined pixel value range.

5. A layered encoder for encoding and decoding a video stream, comprising:

a downsampling unit for reducing the resolution of the video stream;

a base encoder for encoding a lower resolution base stream;

an upconverting unit for decoding and increasing the resolution of the base stream to produce a reconstructed video stream;

a subtractor unit for subtracting the reconstructed video stream from the ginal video stream to produce a residual signal;

a modification unit prior to an enhancement encoder for transforming the residual signal into a signal with a signal level range of a normal input video signal used for video compression;

the enhancement encoder for encoding the resulting residual signal from the modification unit and outputting an enhancement stream.

6. The layered encoder according to claim 5, wherein the modification unit adds a DC-offset value to the residual signal.

7. The layered encoder according to claim 6, wherein the modification unit comprises a clip operation which prevents an output of the modification unit from going below a first predetermined value and above a second predetermined value.

8. The layered encoder according to claim 6, wherein the DC-offset value shifts a majority of pixel values of the residual signal from one end of a predetermined pixel value range to approximately a middle of the predetermined pixel value range.

9. A method for providing spatial scalable compression using adaptive content filtering of a video stream, comprising the steps of:

downsampling the video stream to reduce the resolution of the video stream;

encoding the downsampled video stream to produce a base stream;

decoding and upconverting the base stream to produce a reconstructed video stream;

subtracting the reconstructed video stream from the video stream to produce a residual stream;

transforming the residual signal into a signal with a signal level range of a normal input video signal used for video compression; and

encoding the resulting residual stream and outputting an enhancement stream.

10. The method according to claim 9, wherein the transforming step comprises the steps of:

adding a DC-offset value to the residual signal; and

a clipping operation for preventing an output of the modification unit from going below a first predetermined value and above a second predetermined value.

11. The method according to claim 10, wherein the DC-offset value shifts a majority of pixel values of the residual signal from one end of a predetermined pixel value range to approximately a middle of the predetermined pixel value range.

12. A decoder for decoding compressed video information, comprising:
a base stream decoder for decoding a received base stream;
an upconverting unit for increasing the resolution of the of the decoded base stream;
an enhancement stream decoder for decoding a received enhancement stream;
a modification unit for transforming a signal range of the decoded enhancement stream into a signal range of an original residual signal; and
an addition unit for combining the upconverted decoded base stream and the transformed enhancement stream to produce a video output.

13. A method for decoding compressed video information received in a base stream and an enhancement stream, comprising the steps of:
decoding the base stream;
upconverting the decoded base stream to increase the resolution of the decoded base stream;
decoding the enhancement stream;
transforming a signal range of the decoded enhancement stream into a signal range of an original residual signal; and
combining the upconverted decoded base stream with the transformed enhancement stream to produce a video output.

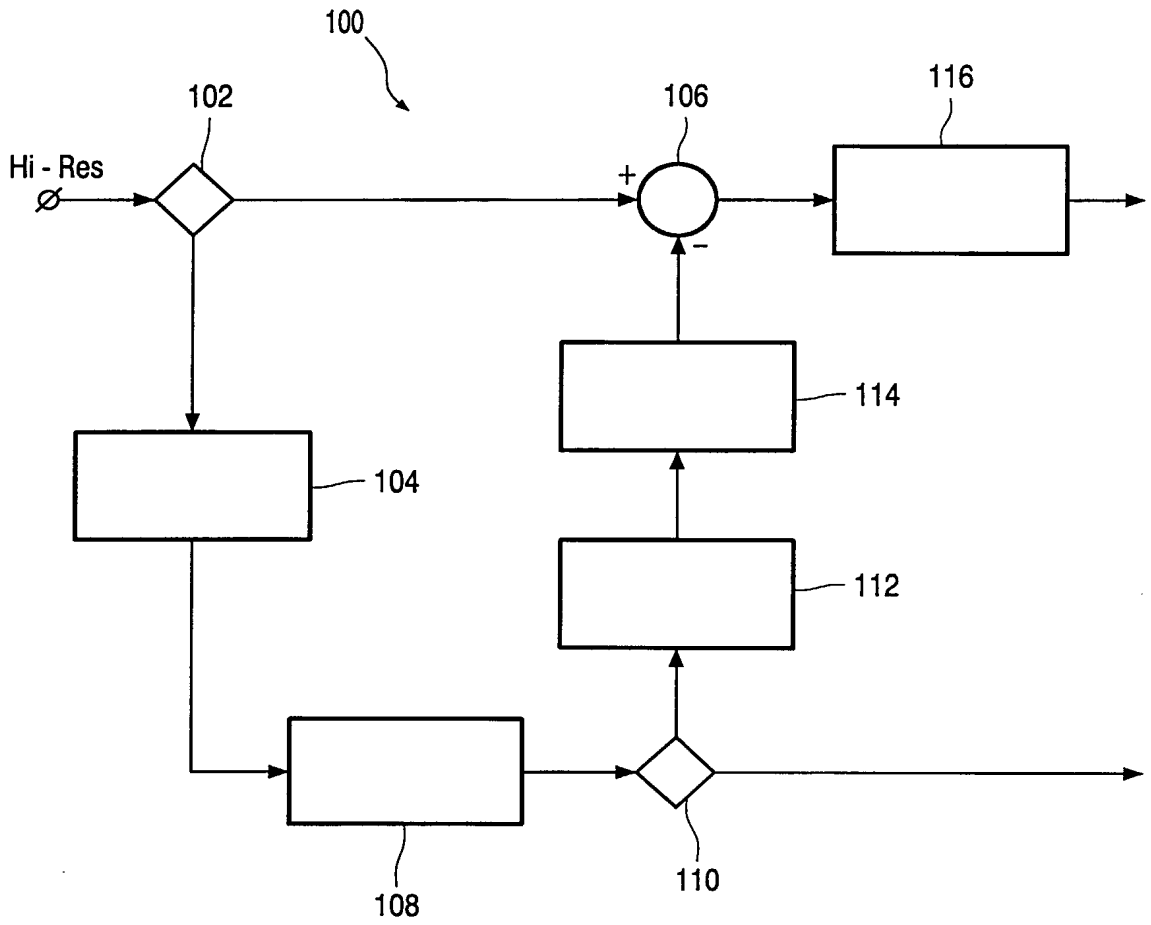


FIG. 1

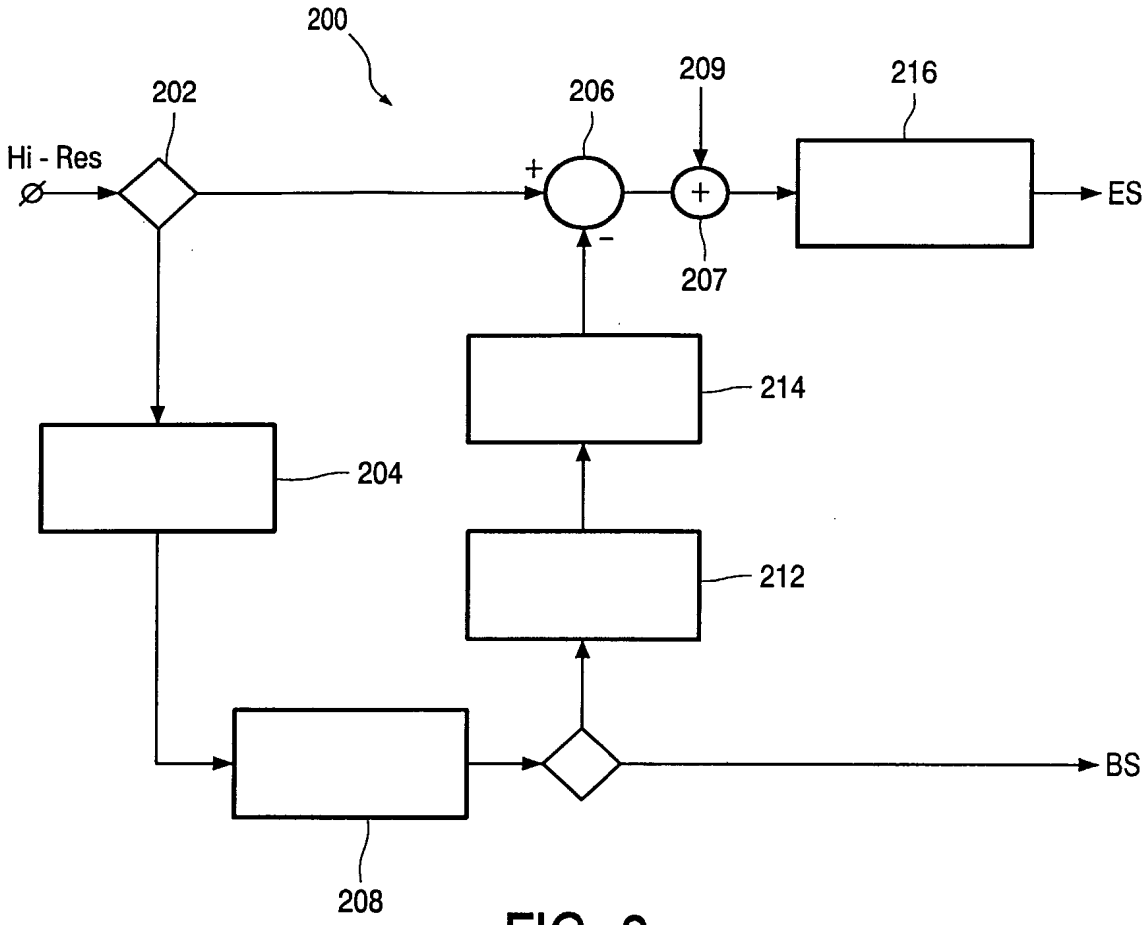

200
202
Hi - Res
206
209
216
+
+
ES
-
207
214
204
212
BS
208

FIG. 2

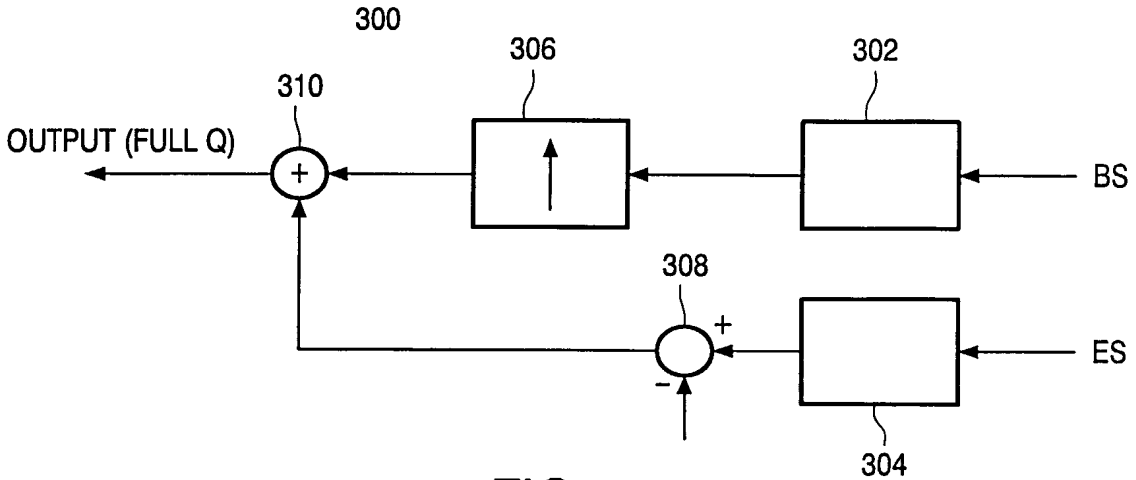

300
306
302
310
OUTPUT (FULL Q)
BS
308
+
ES
-
304

FIG. 3